吴贤铭智能工程学院
SHIEN-MING WU SCHOOL OF INTELLIGENT ENGINEERING

Deep Learning Project on Image Classification

Group No. 8

*** *** ***

2025/06/11

Report Table of Contents

1.Introduction

2. Methods and Implementation

3. Experiment and Result Analysis

4. Innovation point

5. Summary

1. Introduction

- Project Introduction
- Background and Objectives

Research background and Objectives

❑ Research Background

- The development of computer vision: Deep learning promotes the progress of image classification technology
- Application fields: Target recognition, intelligent monitoring, autonomous driving, etc.
- The significance of campus culture: As an important component of campus culture, campus objects possess unique visual features

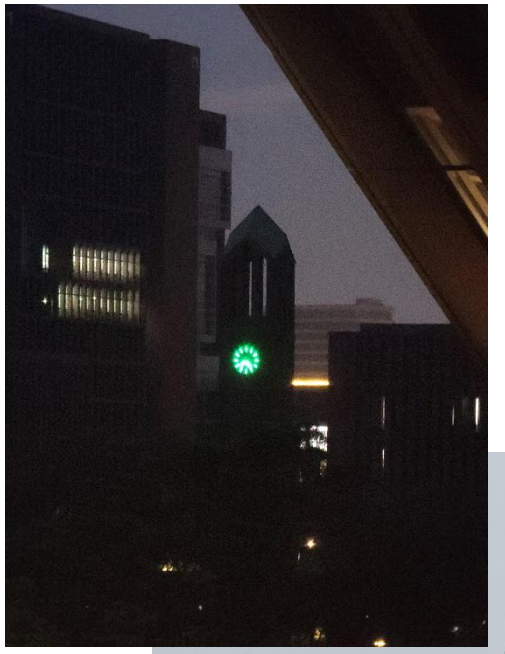
❑ Objective

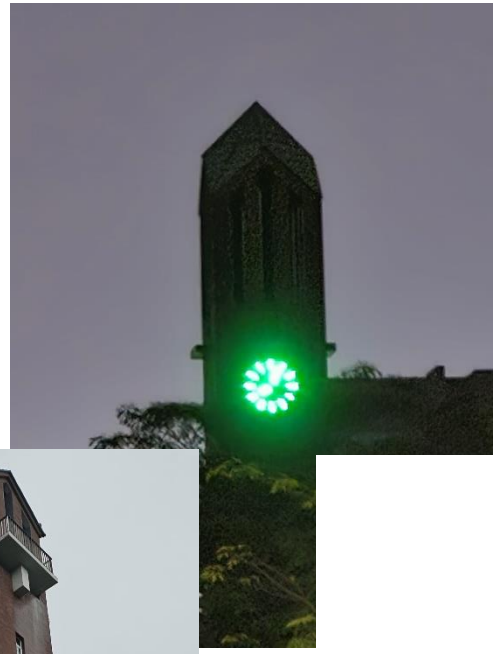
- Build an efficient model: Achieve automatic classification of six types of campus objects
- In-depth understanding of technology: Explore the application of deep learning in image classification through practice
- Innovative approach: Enhance model performance by integrating multiple techniques and strategies

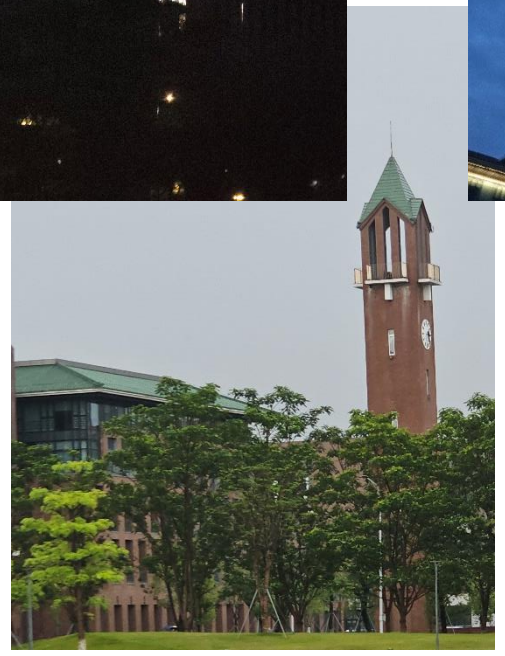
2. Methods and Implementation

- Dataset construction and preprocessing
- Model selection and integration
- Training and Validation

❑ Photo shooting:



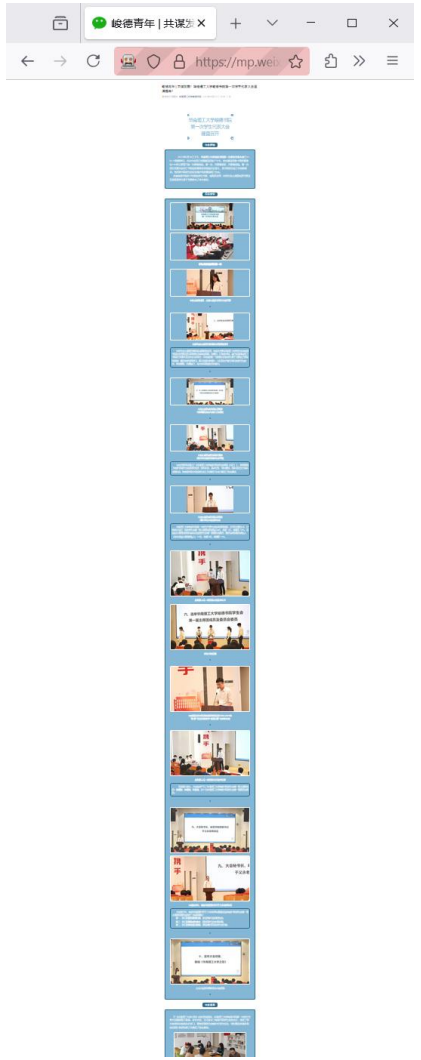




Dataset construction and preprocessing

❑ Python Web Scraping + Reqable Packet Capture:

1) Start web packet capture by Reqable.
2) Use python script to browse all of the result of keyword like “峻德书院/华工鲤鱼君” by Sogou Weixin official account article searching.
3) When browsing each article, use python script to detect if all image loaded.
4) After all articles have been accessed, select all image and export the response body.
5) Filter the image exist Junde mascot.

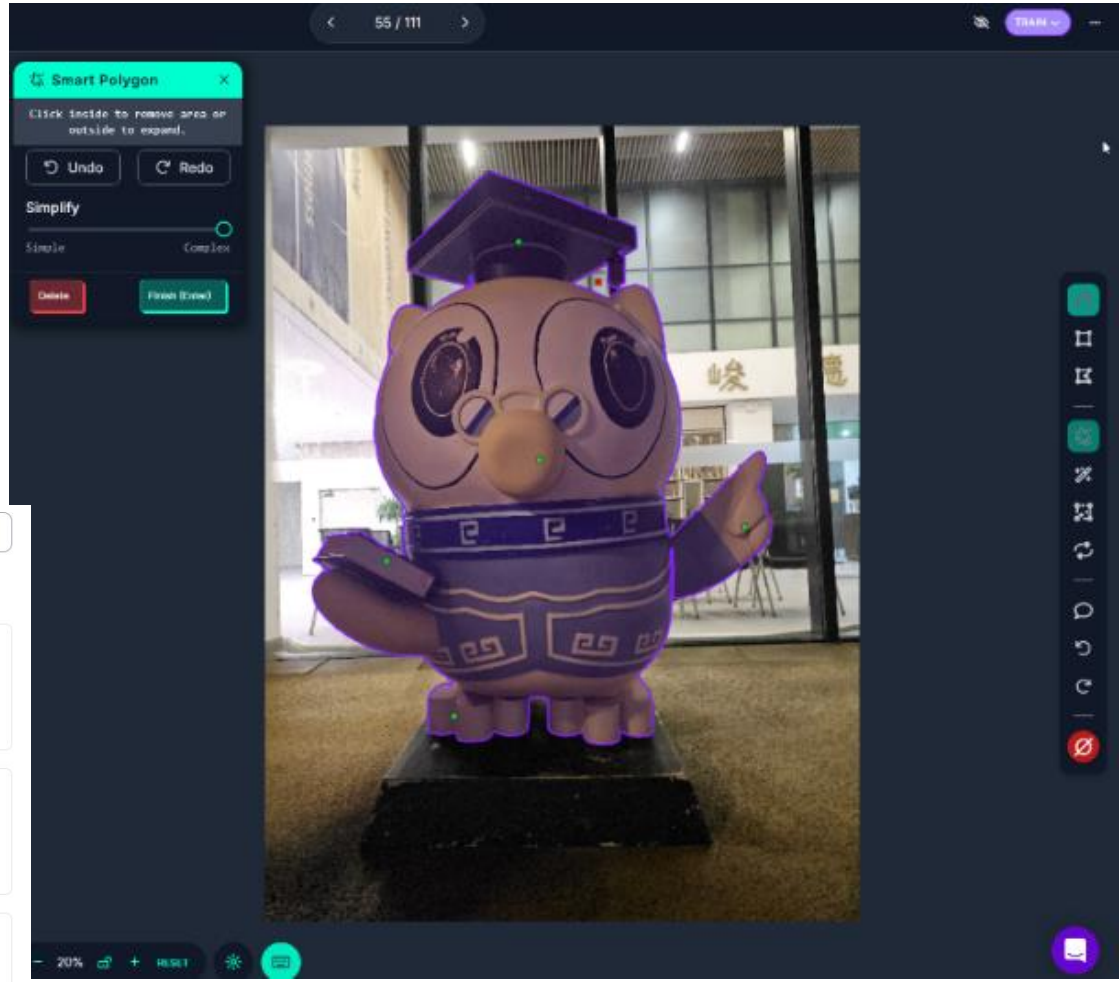


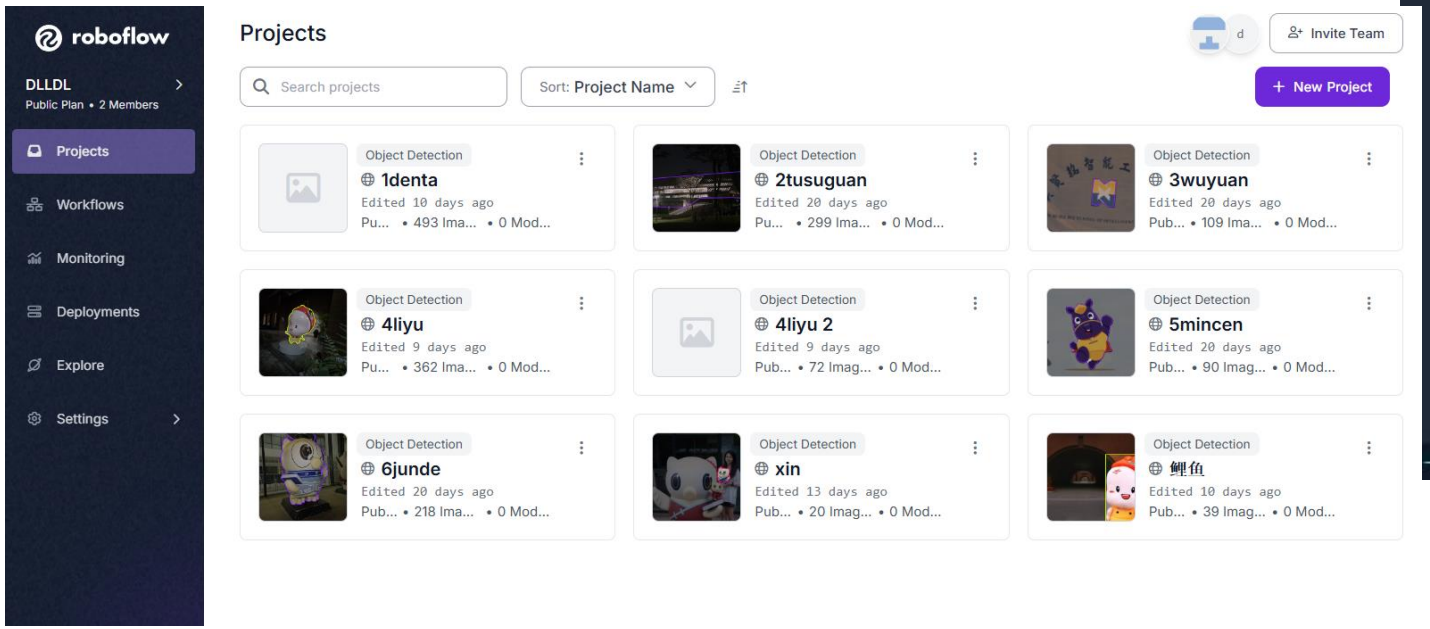
Method

Dataset construction and preprocessing

❑ Dataset construction ：

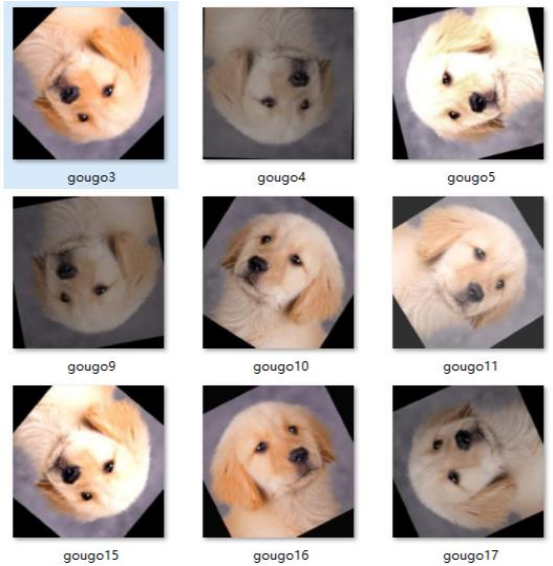
- ➢ Data source
- ➢ Annotation tool
- ➢ Data enhancement

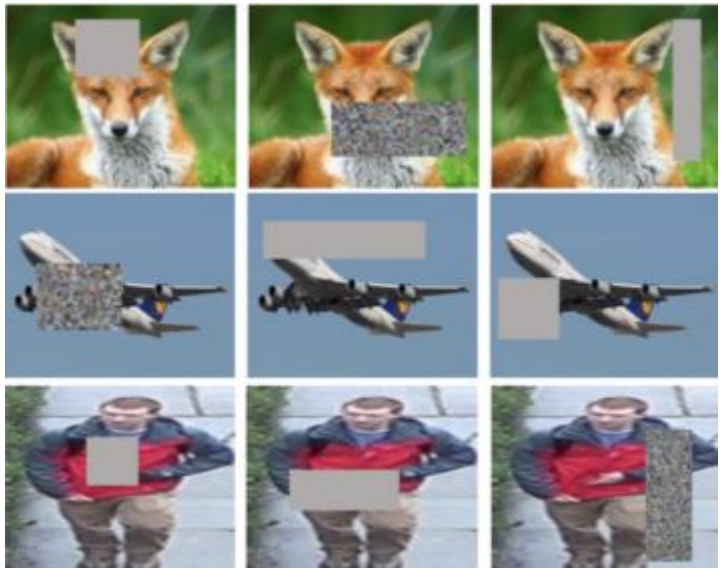




❑ Data preprocessing：

- ➢ Picture adjustment
- ➢ Normalization processing
- ➢ Background removal processing （CustomDataset class）

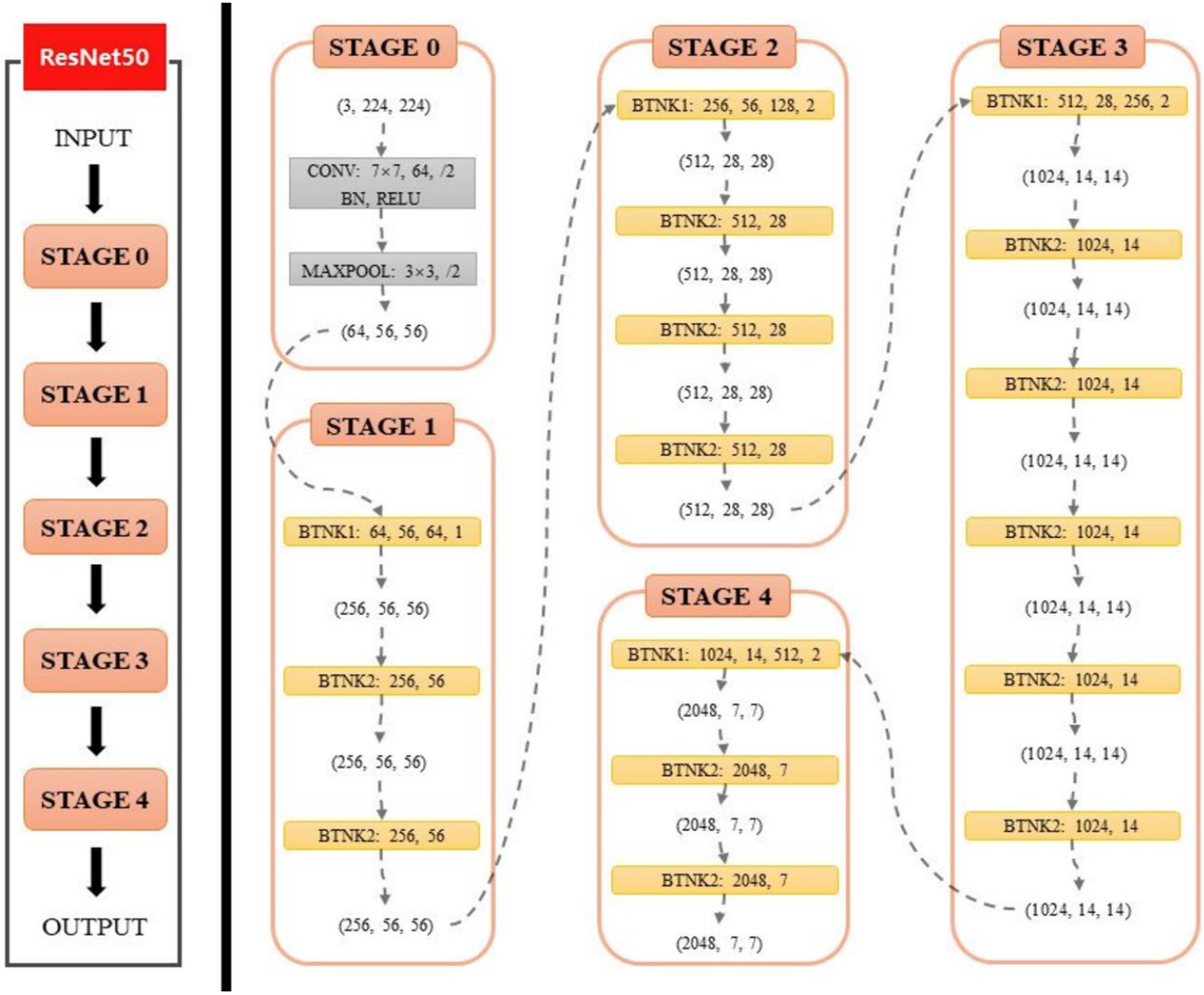




Model selection and integration

❑ ResNet50

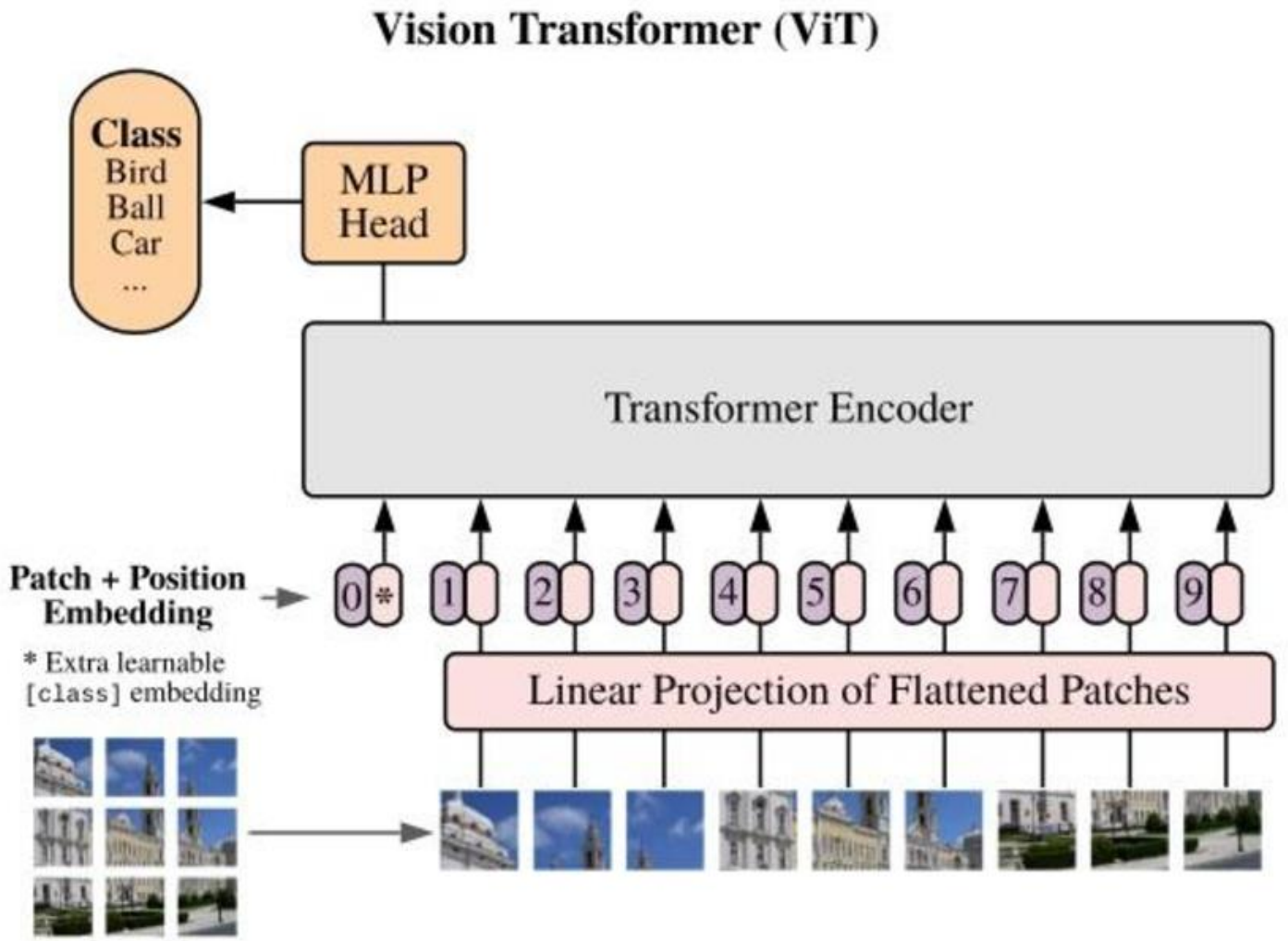
- ➢ Characteristics
- ➢ Advantage



Model selection and integration

❑ ViT

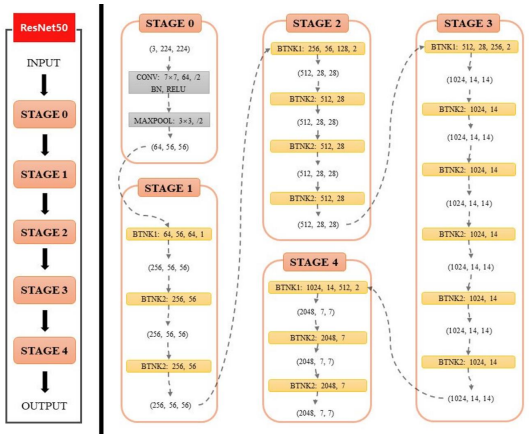
- ➢ Characteristics
- ➢ Advantage

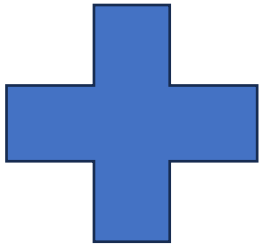


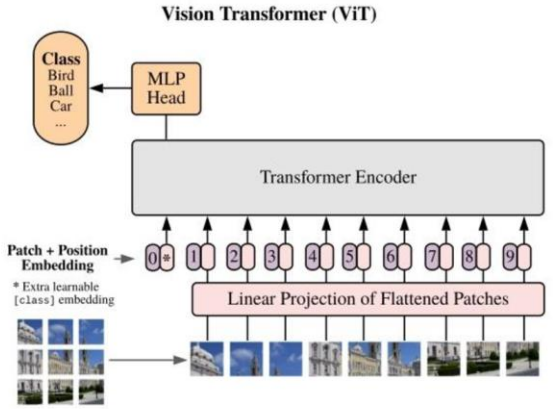
Model selection and integration

❑ Ensemble model joint training

- ➢ reason
- ➢ advantage







ResNet50+ViT Ensemble model

Training and Validation

❑ Dataset partitioning

- ➢ Allocation

❑ Training process

- ➢ Loss function
- ➢ Optimizer
- ➢ Learning rate scheduling
- ➢ Training stage
- ➢ Verification stage

❑ Model update

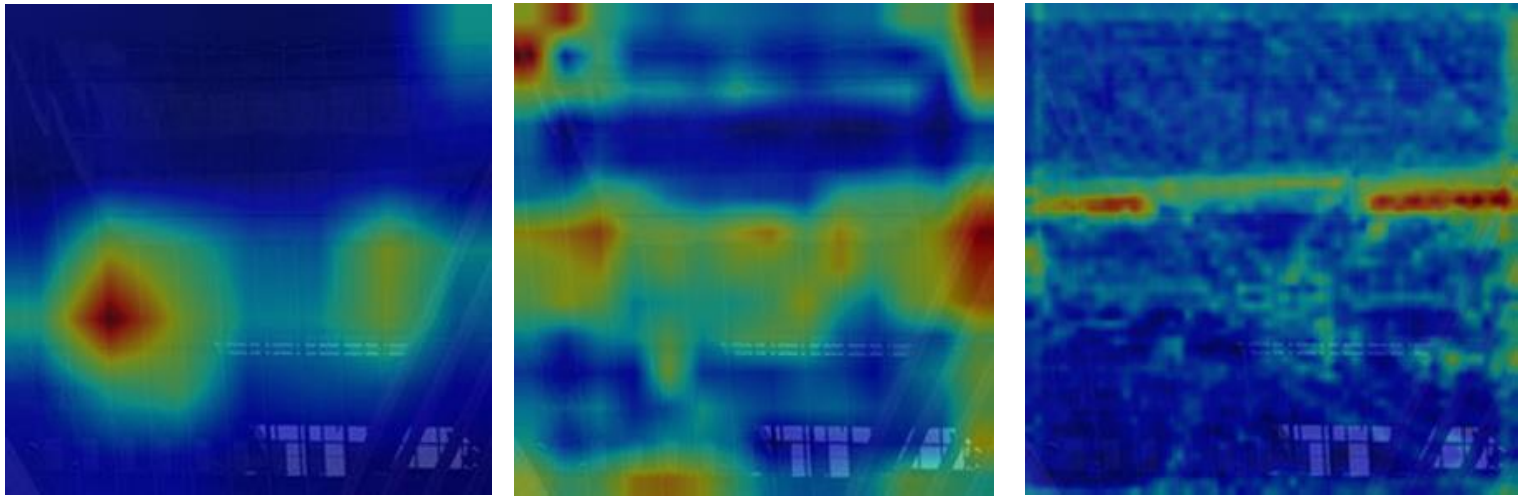
- ➢ Best model preservation
- ➢ Performance monitoring

3. Experiment and Result Analysis

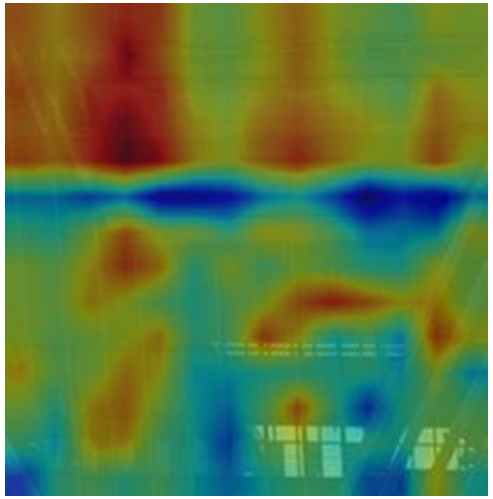
- Experimental verification
- Performance analysis
- Model performance

Feature extraction feature verification

- Experiment One: Intermediate Decision-making



The intermediate feature map of the library by resnet50

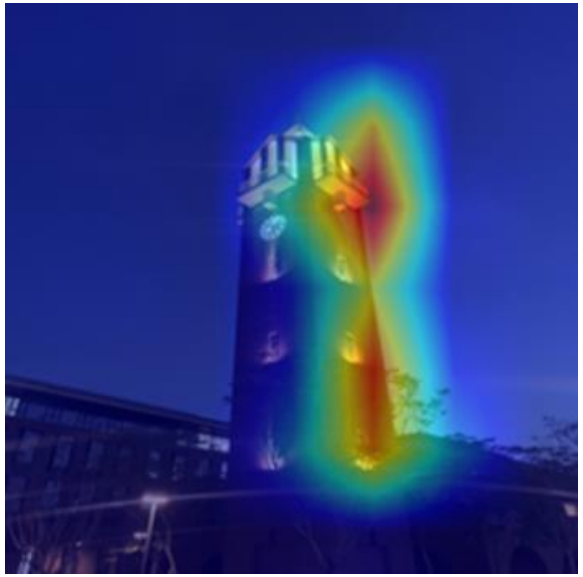


The intermediate feature map of the library by ViT

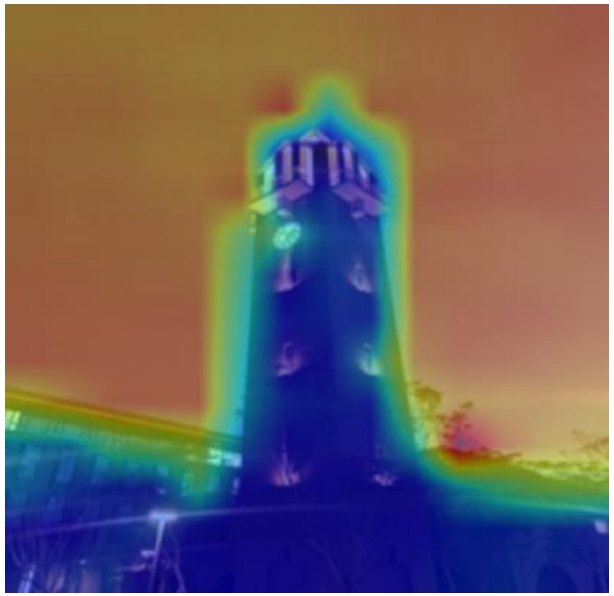
Result

Feature extraction feature verification

- Experiment Two: Final Decision



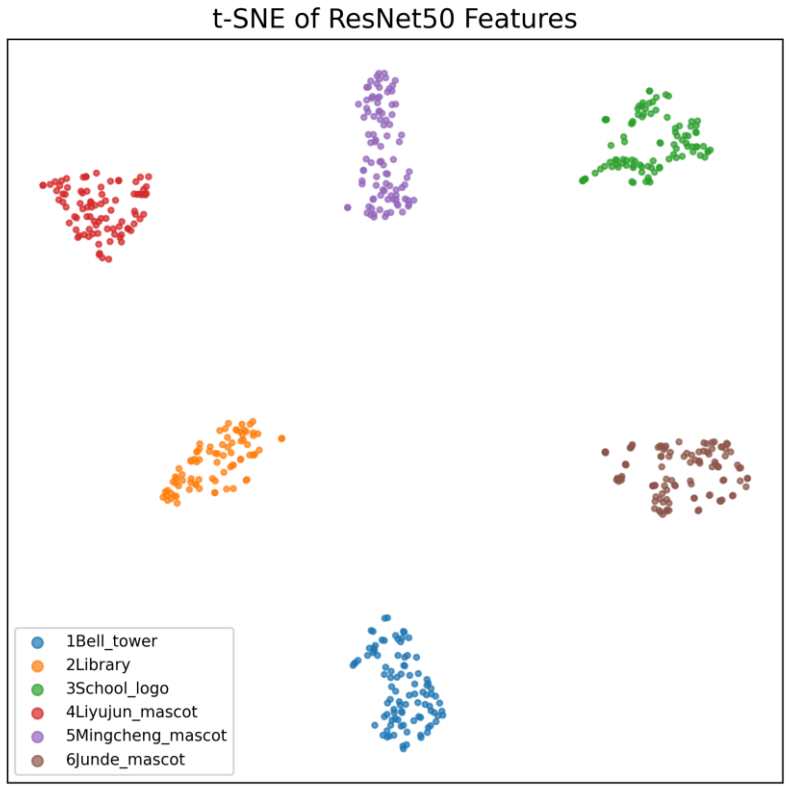
The final activation diagram of resnet50 for the clock tower

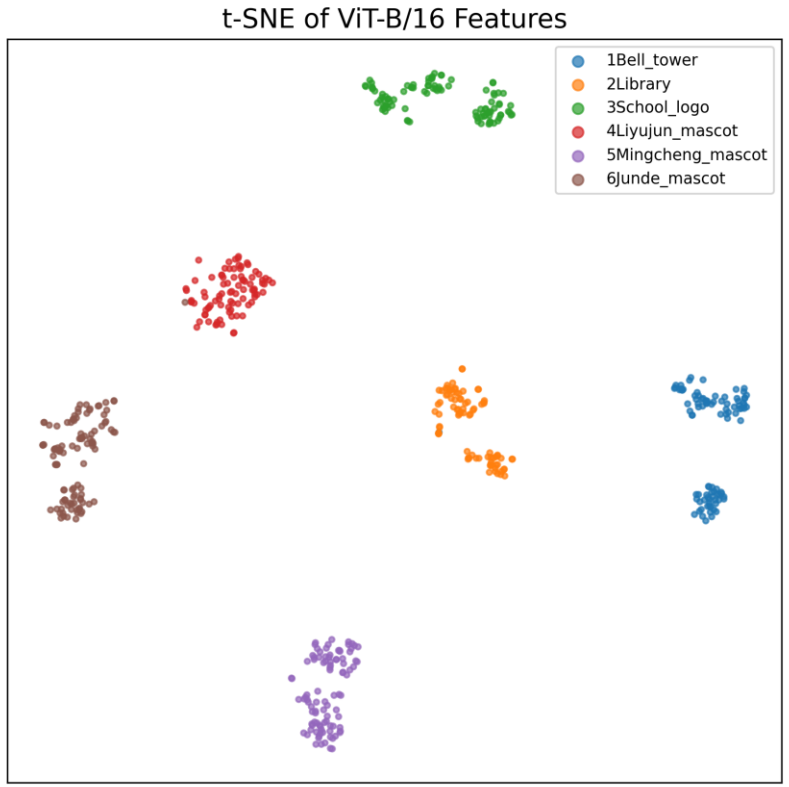


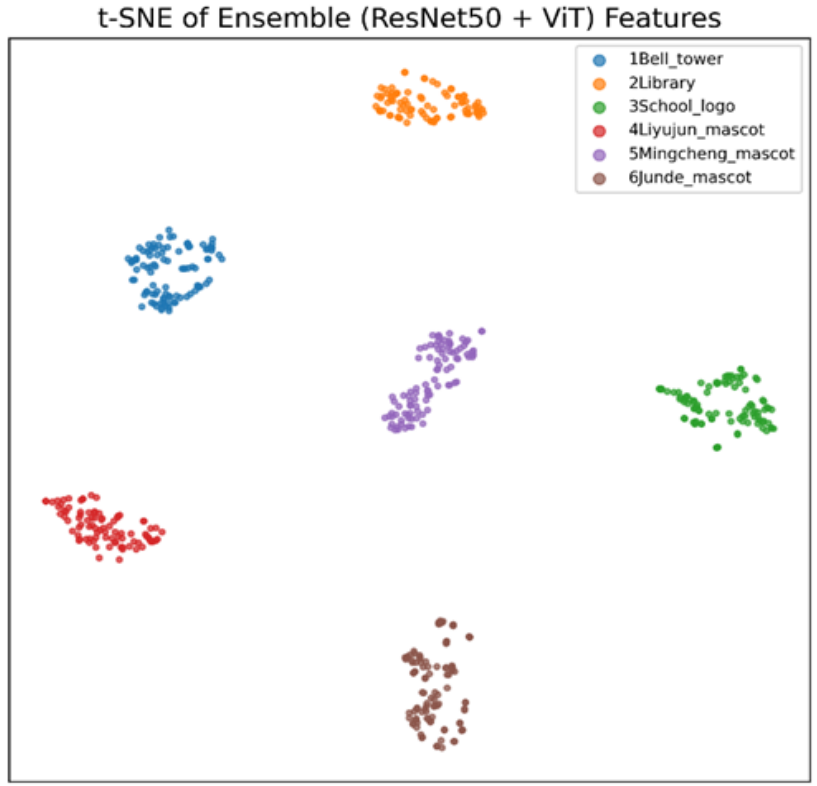
The final activation diagram of the clock tower by ViT

Result Ablation experiment

- Experiment 3: Cluster Analysis of Mapping High-dimensional Features to a two-dimensional space



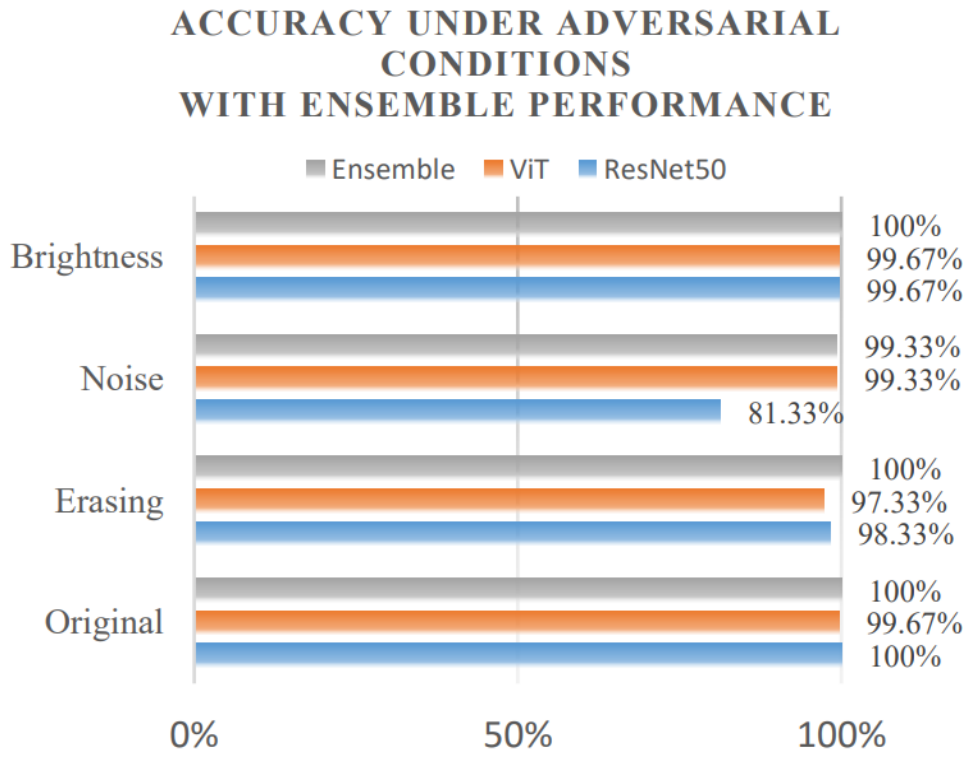




Result Ablation experiment

- Experiment Four: Robustness Analysis Experiment

Mode	Resnet50 (%)	ViT (%)	Ensemble (%)
Original	100	99.67	100
Erasing	98.33	97.33	100
Noise	81.33	99.33	99.33
Brightness	99.67	99.67	100



❑ The advantages of model performance

- ResNet50 single model: Test set accuracy rate 86.67%
- ViT single model: The accuracy of the test set is 83.33%
- Ensemble model +TTA: The accuracy rate of the test set is 90%
- Comparative analysis: The ensemble model is significantly superior to the single model, and the ensemble model has more significant advantages when the data volume of the test set increases, which is attributed to its excellent robustness

4. Innovation Point

- Background removing
- Data enhancement
- Cross-architecture integration
- TTA Strategy
- One-click training
- Optimize data loading

5. Summary

- Key conclusion
- Model decision-making mechanism
- Future optimization path

Conclusion Project Summary and Future Outlook

❑ Key conclusion

- Background removal processing reduces background interference
- Cross-architecture integration enhances performance
- TTA enhances the robustness of the model

❑ Model decision-making mechanism

- Visualize the activation map and feature map
- Cluster analysis verifies the feature extraction ability

❑ Future work

- Introduce more diverse data augmentation technologies
- Explore deeper model fusion methods
- Make use of more efficient training techniques
- The style of the data set has become more diverse

Co-Work Individual Workload of Students

❑ Student A

- Conducted preliminary literature review; Developed data preprocessing scripts (including but not limited to data augmentation and background removal using masks); Built, selected and trained models; Designed and conducted experiments; Making 9 charts for Report; Assisted in refining the report and presentation slides; Took charge of the presentation and Q&A session.

❑ Student B

- Coded a Web Scrap for Sogou; Took, collected and annotated over 1600 photos; Translated, revised and added content to Report; Made 1 chart and 6 formulas for Report; Proofread and added content to PowerPoint; Edited and added content to speech script.

❑ Student C

- Took, collected and annotated approximately 300 images; Drafted the initial Chinese version of the report, created presentation slides and the speech script; Making 1 chart for Report; Presented the content using the slides.



吴贤铭智能工程学院
SHIEN-MING WU SCHOOL OF
INTELLIGENT ENGINEERING

Thanks